

Journal of Surgery
[Jurnalul de Chirurgie]

Original paper Open Access

Applications using artificial intelligence in predicting the risk of postoperative pancreatic fistula using preoperative CT images

Marius Lucian Savin, Liliana Gheorghe, Corina Lupașcu Ursulescu, Dragoș Negru, Ana Maria Trofin, Mihai Zabara, Vlad Nuțu, Ramona Cadar, Mihaela Blaj, Oana Lovin, Felicia Crumpei, Cristian Lupașcu
Clinica I Chirurgie, Universitatea de Medicină și Farmacie ”Grigore T. Popa”, Spitalul Clinic de Urgențe ”Sf. Spiridon” Iași

Abstract.Background: Despite the progress in surgical technique, postoperative pancreatic fistula (POPF) remains the main cause of morbidity and mortality after pancreatectomy. Estimating the risk of fistula is of paramount importance, guiding different strategies to mitigate the global morbidity of the patients after surgery. Existing risk scores take into account pre-, intra- and immediate postoperative parameters. **Methods.** We investigated the accuracy of convolutional neural networks (CNN) in the estimation of the risk of POPF using solely preoperative CT images, using 4 classic methods. **Results.**The VGG16 model showed a good predictive power both in training and internal testing datasets. **Conclusion.** Machine learning methods have the potential to further aid the surgeon in assessing the risk of POPF due to the automation advantage, but require further investigation, validation as compared to classical risk scores.

Keywords: artificial intelligence in surgery; postoperative pancreatic fistula, CT

Introducere

Pancreatectomia reprezintă tratamentul optim pentru majoritatea tumorilor benigne și maligne ale pancreasului. Rata de mortalitate după duodeno-pancreatectomia cefalică (DPC), cel mai frecvent tip de pancreatectomie, indiferent de tehnica utilizată (cu sau fără prezervare de pilor), deși în scădere în ultimele decenii [1,2], rămâne semnificativ crescută [3–5], de 10-28% [6]. Fistula pancreatică post-operatorie după DPC poate favoriza apariția de hemoragie intra-abdominală și peritonită severă [7,8], rămâne cauza principală de morbiditate post-operatorie, costuri de spitalizare crescute sau chiar moarte (5–8).

Rezultă necesitatea identificării pacienților cu risc crescut. Studiile precedente au recunoscut următorii factori imagistici predictivi: pancreasul moale, lipomatoza pancreatică, calibrul ductului pancreatic, obezitatea [12–14].

Obiectivului studiului actual este validarea unor modele de învățare automată care să prezică cu acuratețe înaltă fistula pancreatică *semnificativă clinic* (grad B și C) după pancreatectomie, folosind imaginile obținute în urma examenului computer tomografic (CT).

Învățarea profundă bazată pe imagini, poate prezice cu înaltă precizie fistula post-operatorie după pancreatectomie, folosind imaginile tomografiei computerizate (CT) și are potențial ridicat pentru aplicații generalizate în interpretarea imaginilor biomedicale și luarea deciziilor medicale.

Repere:

- S-a dezvoltat un sistem de inteligență artificială care

Corresponding author: Zabara Mihai Lucian ,”Grigore T. Popa”University of Medicine and Pharmacy Iași. E-mail zabara_mihai@yahoo.com

Received: May 30, 2022; **Accepted June** 29, 2022; **Published September** 30, 2022

Citation: Savin M.L., Gheorghe Liliana, Lupașcu Ursulescu Corina, Negru D., Trofin Ana Maria, Zabara M., Nuțu V., Cadar Ramona, Blaj Mihaela, Lovin Oana, Crumpei Felicia , Lupașcu Cr. **Applications using artificial intelligence in predicting the risk of postoperative pancreatic fistula using preoperative CT images.** Journal of Surgery [Jurnalul de chirurgie]. 2022; 18(3): 215 - 223, [Article in Romanian]. DOI: 10.22551/JSURG.2022.03.05

Copyright: © 2022. Savin M.L... This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited

- utilizează tehnici de învățare prin transfer
- S-au clasificat eficient imaginile CT pentru fistula post-operatorie după pancreatectomie
- Această procedură are potențial pentru o aplicare generalizată cu impact mare în imagistica biomedicală.

Inteligența artificială (AI) are potențialul de a revoluționa diagnosticul și managementul maladiilor prin efectuarea unor clasificări dificile pentru experții umani și prin revizuirea rapidă a unor cantități imense de imagini. În ciuda potențialului său, interpretabilitatea clinică și pregătirea fezabilă a AI rămân o provocare.

Abordarea algoritmică tradițională a analizei imaginilor pentru clasificare s-a bazat anterior pe (1) segmentarea obiectelor realizată manual, urmată de (2) identificarea fiecărui obiect segmentat folosind clasificatoare statistice sau clasificatoare neuronale superficiale de tip învățare automată concepute special pentru fiecare clasă de obiecte, și în final (3) clasificarea imaginilor [1]. Crearea și perfecționarea mai multor clasificatori a necesitat mulți oameni calificați și mult timp și a fost costisitor din punct de vedere al calculului computerizat[15,16]

Dezvoltarea straturilor convoluționale de rețea neuronală a permis câștiguri semnificative în capacitatea de a clasifica imagini și de a detecta obiecte dintr-o imagine [17,18]. Este vorba despre mai multe straturi de procesare la care se aplică filtre de analiză sau convoluții. Reprezentarea abstractizată a imaginilor în cadrul fiecărui strat este construită transformând sistematic mai multe filtre peste imagine, producând o hartă de caracteristici care este utilizată ca intrare pentru următorul strat. Această arhitectură permite procesarea imaginilor sub formă de pixeli ca intrare și pentru a da clasificarea dorită ca ieșire. Abordarea imagine-spre-clasificare într-un clasificator înlocuiește mai multe etape ale metodelor anterioare de analiză a imaginii.

O metodă de abordare a lipsei datelor dintr-un anumit domeniu este de a utiliza datele dintr-un domeniu similar, o tehnică cunoscută sub numele de învățare prin transfer. Învățarea prin transfer s-a dovedit a fi o tehnică extrem de eficientă, în special atunci când se confruntă cu domenii în care datele sunt limitate [19–21]. În loc să antrenăm o rețea complet goală, folosim o abordare de tip feed-forward pentru a seta ponderile la nivelele inferioare deja optimizate pentru a recunoaște structurile găsite în imagini în general și pentru a reantrena ponderile nivelurilor superioare cu back-propagation, modelul poate recunoaște caracteristicile distinctive ale unei anumite categorii de imagini, cum ar fi imaginile ochiului, mult mai rapid și cu exemple mult mai puține de antrenare și cu mai puțină putere de calcul (Figura 1).

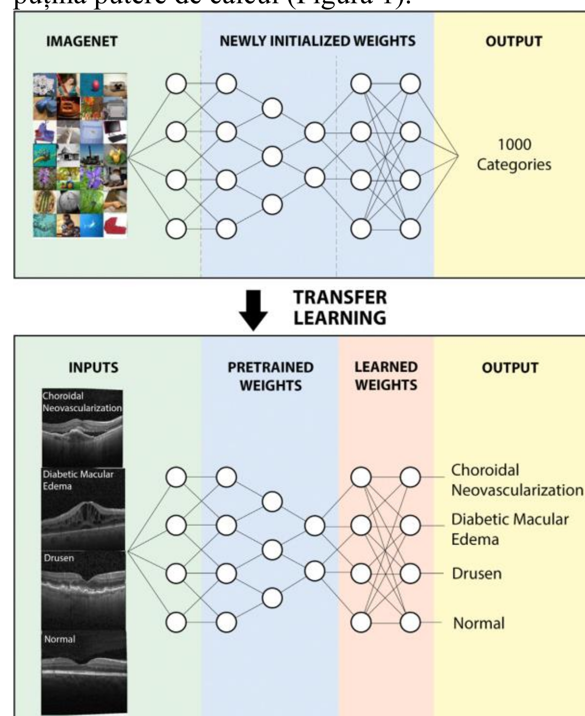


Figura 1. Schema unei rețele neuronale convoluționale - Schemă care ilustrează modul în care o rețea neuronală convoluțională antrenată pe setul de date ImageNet din 1.000 de categorii poate fi adaptată pentru a crește în mod semnificativ precizia și pentru a scurta durata de antrenare a unei rețele antrenate pe un nou set de date cu imagini CT. Straturile conectate local (convoluțional) sunt înghețate și transferate într-o nouă rețea, în timp ce straturile finale, complet conectate sunt recreate și reantrenate de la inițializarea aleatorie în partea de sus a straturilor transferate.

Material și metodă.

Examenul CT preoperator a fost efectuat între 2 și 58 zile preoperator (mediana =21). Pentru 28 din cei 33 pacienți, scanarea post-operatorie precoce a fost efectuată cu un computer tomograf (CT) cu 16 coroane de detectori Somatom Sensation 16 (Siemens Medical Solutions, Erlangen, Germany). Câmpul de scanare s-a extins de la 2 cm superior de diafragm până la nivelul simfizei pubiene.

Am împărțit studiul nostru pe 4 subgrupe, funcție de secvențele achiziționate (nativ și venos) și de tipul de fistulă pancreatică (semnificativă și nesemnificativă) și le-am clasificat în funcție de o stare normală :

- nativ – semnificativ/normal : train – 200/300, test – 109/113, validate – 8/8
- nativ – nesemnificativ/normal : train – 300/300, test - 140/113, validate – 8/8
- venos – semnificativ/normal : train - 300/370, test – 92/103, validate – 8/8
- venos – nesemnificativ/normal: train – 400/370, test – 109/103, validate – 8/8

Învățarea prin transfer

Acordul-fin al modelului nostru

Implementarea de acord/reglare fină constă în trunchierea ultimului strat (stratul softmax) al rețelei pre-instruite și înlocuirea lui cu un nou strat softmax care este relevant pentru analiza noastră.

Au fost evaluate mai multe modele de rețea neuronală, care sunt descrise sumar în cele ce urmează.

1) MobileNet

MobileNet este o arhitectură care este mai potrivită pentru aplicațiile de vizionare bazate pe dispozitive mobile și încorporate, în care lipsește puterea de calcul. Această arhitectură a fost propusă de Google. Implementarea aleasă este detaliată în Anexa 1.

2)VGG16

Cercetătorii de la Oxford Visual Geometry Group sau VGG pe scurt, au dezvoltat rețeaua VGG, care se caracterizează prin simplitatea sa, folosind doar 3 x 3 straturi convoluționale stivuite una peste alta, în profunzime crescândă. Reducerea dimensiunii volumului se face prin max pooling. La final, două straturi complet conectate, fiecare cu 4.096 noduri, sunt apoi urmate de un strat softmax.

Pooling-ul se realizează prin straturi max pooling, care urmează unele dintre straturile de convoluție. Nu toate straturile de convoluție sunt urmate de max pooling. Max pooling se realizează pe o fereastră de 2 x 2 pixeli, cu un pas/stride de 2. Activarea ReLU este utilizată în fiecare strat ascuns. Numărul filtrelor crește cu adâncimea în majoritatea variantelor VGG (Figura 2).

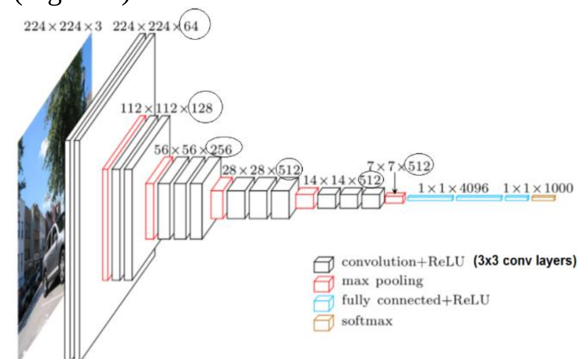


Figura 2 Arhitectura cu 16 straturi VGG-16

În mod similar cu mobilenet, pașii sunt efectuați cu puține modificări.
Implementarea acestuia este detaliată în Anexa 2.

3) Model CNN cu 5-Straturi ascunse

Aplicăm 5 CNN (strat convoluțional) cu nucleu 3 * 3 și activări „relu”.
Implementarea acestuia este detaliată în Anexa 3.

4)Model VGG Customizat

Implementarea este detaliată în Anexa 4.

Rezultate:

1) Model 1 MobileNet

Acuratețea testului este de 85.135 %

Matricea de Confuzie

Confusion Matrix
[[80 33]
 [0 109]]

Classification Report

	precision	recall	f1-score	support
normal	1.00	0.71	0.83	113
semnificativ	0.77	1.00	0.87	109
accuracy			0.85	222
macro avg	0.88	0.85	0.85	222
weighted avg	0.89	0.85	0.85	222

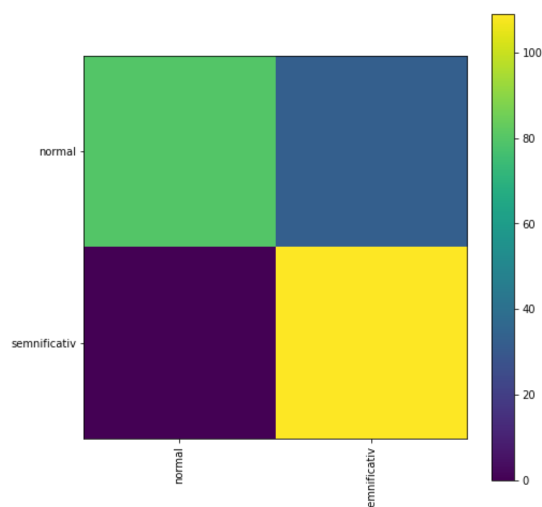


Fig.3 Imaginile care reprezintă fistula semnificativă pot fi identificate cu mai multă ușurință în raport cu cele normale.

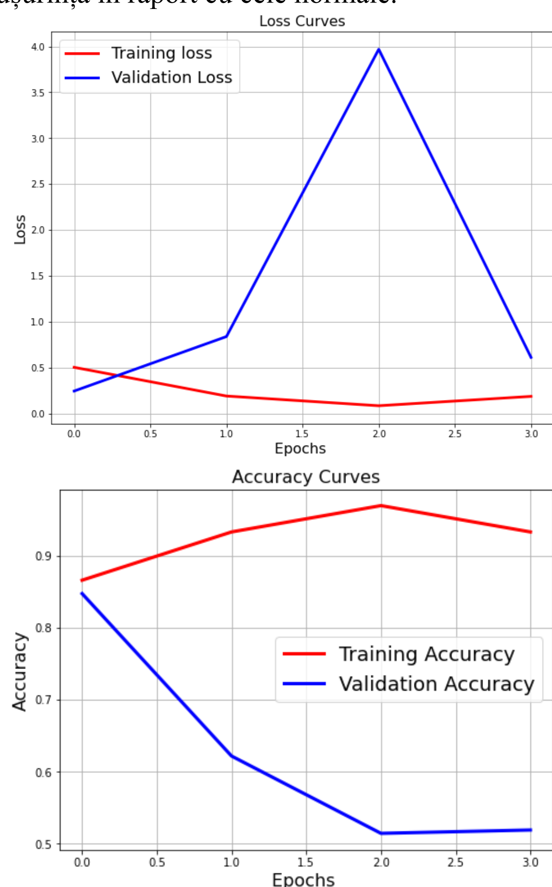


Fig.4 Curbele accuracy și loss indică un overfitting al modelului. În continuare se observă un rezultat mulțumitor pentru curba ROC și un AUC de 0.854.

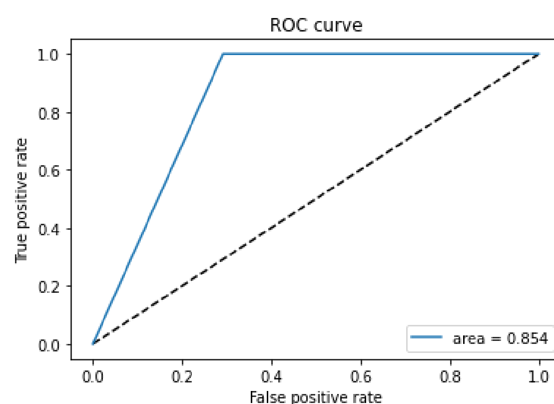


Fig. 5 Rezultatul a putut fi îmbunătățit suplimentar lucrând cu alte modele.

2) Model 2 VGG16

Acuratețea testului este de: 100.00.

Matricea de confuzie

Confusion Matrix

[[113 0]

[0 109]]

Classification Report

	precision	recall	f1-score	support
normal	1.00	1.00	1.00	113
semnificativ	1.00	1.00	1.00	109
accuracy			1.00	222
macro avg	1.00	1.00	1.00	222
weighted avg	1.00	1.00	1.00	222

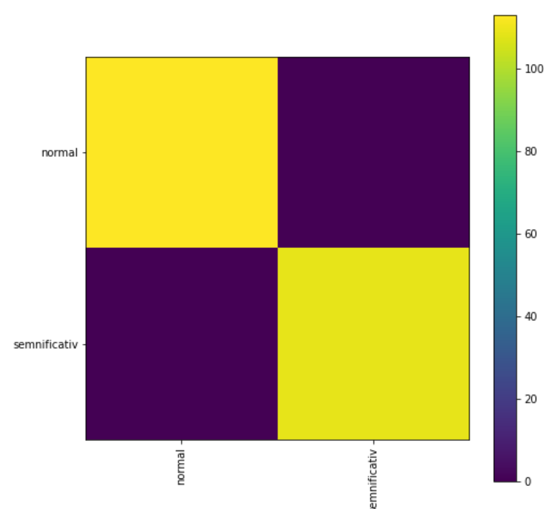


Fig. 6 În cazul acestui model imaginile normale și cele care reprezintă fistula semnificativă pot fi identificate cu ușurință și cu acuratețe maximă.

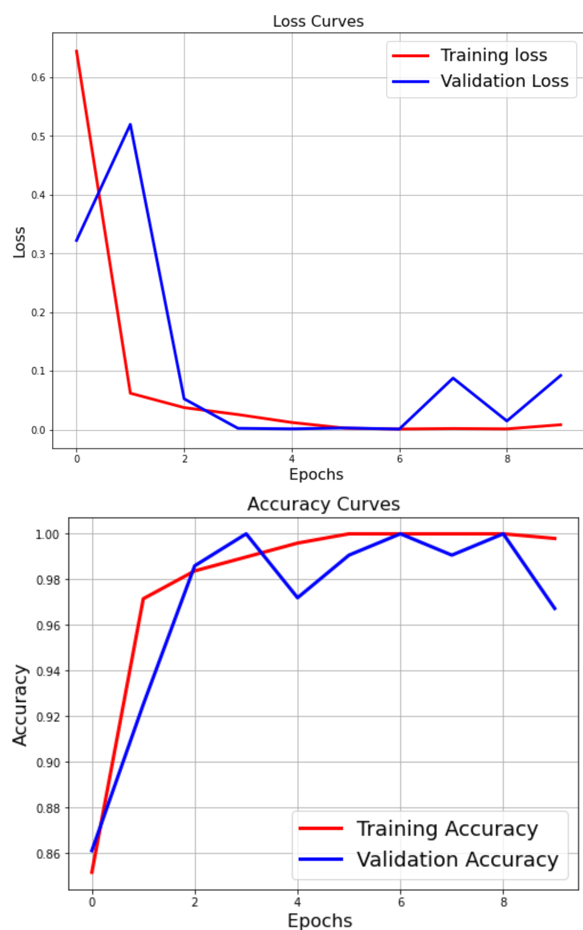


Fig. 7 Curbele accuracy și loss nu indică semne de overfitting al acestui model, ceea ce înseamnă că el se va comporta bine atunci când va fi alimentat cu imagini noi, pe care nu le-a "văzut" niciodată. În continuare se observă un rezultat excelent și pentru curba ROC și un AUC maxim de 1.0.

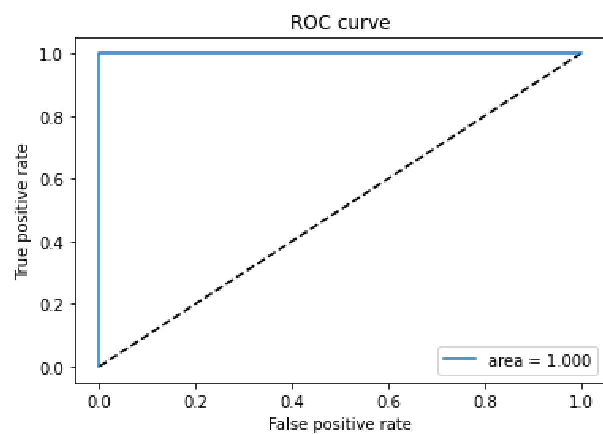


Fig. 8 Curba ROC si AUC

3) Model CNN cu 5-Straturi ascunse

Acuratețea Testului

Test result: 68.018 loss: 3.543

Matricea de confuzie

Confusion Matrix
[[113 0]
[71 38]]

Classification Report

	precision	recall	f1-score	support
normal	0.61	1.00	0.76	113
semnificativ	1.00	0.35	0.52	109
accuracy			0.68	222
macro avg	0.81	0.67	0.64	222
weighted avg	0.80	0.68	0.64	222

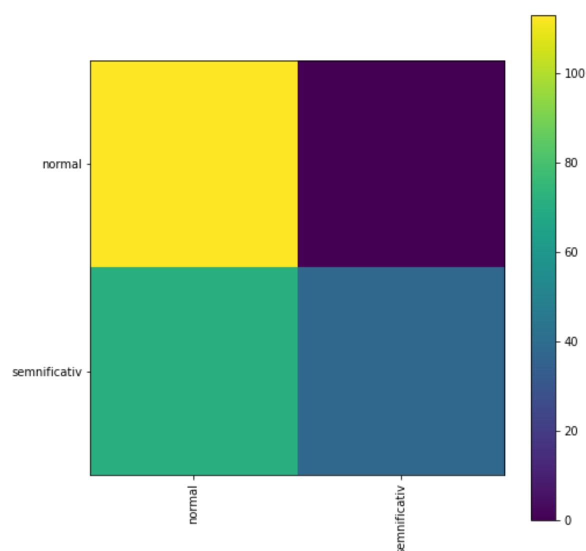


Fig. 9 gImaginile care reprezintă starea normală pot fi identificate cu mai multă ușurință în raport cu cele care reprezintă fistula semnificativă.

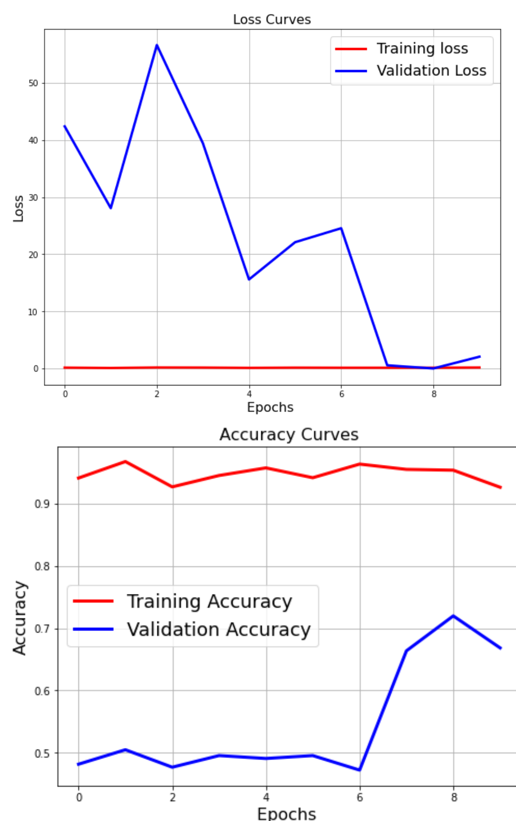


Fig. 10 Curbele accuracy și loss indică un puternic overfitting al modelului. În continuare se observă un rezultat slab pentru curba ROC și un AUC de 0.674.

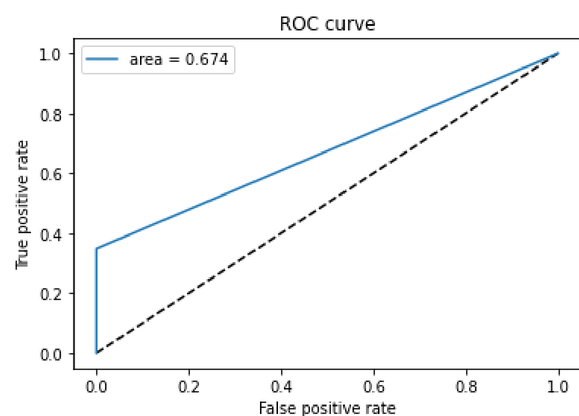


Fig. 11 Curba ROC si AUC de 0.674

4) Model VGG Customizat

Matricea de confuzie

Confusion Matrix
[[0 113]
 [0 109]]

Classification Report

	precision	recall	f1-score	support
normal	0.00	0.00	0.00	113
semnificativ	0.49	1.00	0.66	109
accuracy			0.49	222
macro avg	0.25	0.50	0.33	222
weighted avg	0.24	0.49	0.32	222

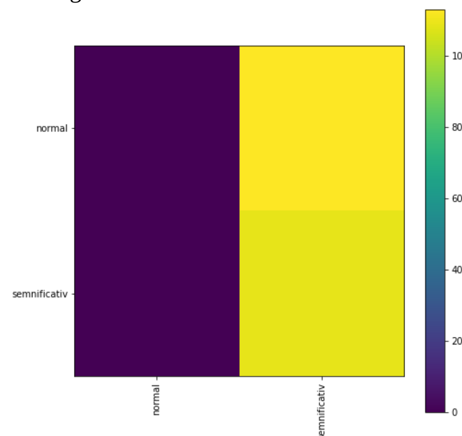


Fig. 12 Modelul a prezis corect toate cazurile de fistula semnificativă, dar a greşit toate cazurile normale.

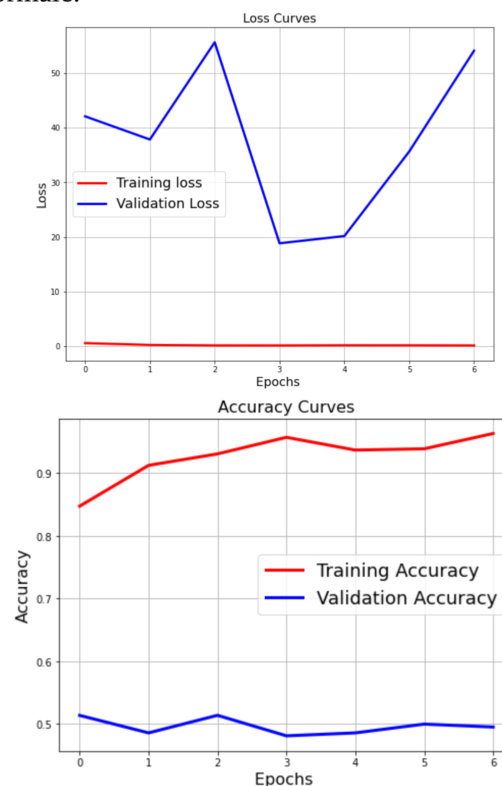


Fig. 13 Curbele accuracy și loss indică faptul că modelul nu a reușit să învețe din imagini. În continuare se observă un rezultat nul pentru curba ROC și prin urmare un AUC de 0.500.

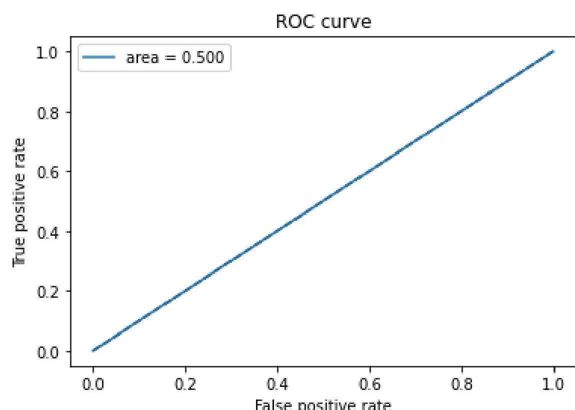


Fig. 14 Curba ROC si AUC de 0.500

Discuții

Au fost formulate multiple scoruri ce cuantifică riscul de fistulă pancreatică semnificativă clinic.

Scorul formulat de Callery în 2013 (fistula risk score FRS) folosește parametri intraoperatori, anume prezența pancreasului moale, diametrul ductului Wirsung, prezența de patologie alta decât adenocarcinom și pancreatită cronică, precum și hemoragia intra-operatorie [88].

Mungroop a formulat un scor alternativ (a-FRS) bazat pe parametri mai simplu de estimat, anume textura pancreasului, diametrul ductului Wirsung și indicele de masă corporală [54]. Același autor a creat un scor updatat, creat în principal pentru chirurgia pancreatică minim invazivă, dar validat ulterior pentru chirurgia descrisă, care include și sexul masculin ca factor adițional [89].

Kantor a formulat în 2017 un scor similar, de asemenea validat extern, ce ia în considerație ca parametru sexul, indicele de masă corporală, nivelul bilirubinei preoperatorii, diametrul ductului Wirsung și textura pancreasului.

O meta-analiză realizată de Schuh în 2021 pe baza a 108 studii relevante a identificat o clasificare în patru tipuri a modificărilor pancreatice a factorilor de risc ce țin de pancreas [96]. Tipul A include pancreasul non-moale cu diametrul ductului Wirsung >3 mm. Tipul B se referă la un pancreas non-moale cu diametru <3 mm. Tipul C se referă la un pancreas moale cu diametrul ductului Wirsung >3 mm. Tipul D include pacienți cu pancreas moale și duct Wirsung <3 mm. Riscul de fistulă crește substanțial în tipul D (23.2%) față de tipul A (3.5%).

Recent, s-au creat nomograme cu specificitate mare bazate pe probele paraclinice obținute în ziua 1 postoperatorie [92].

Impactul real al scorurilor de predicție începând cu perioada preoperatorie, cu consilierea bazată pe dovezi, continuă intraoperator și post-operator, cu ajustarea evaluării și managementului pentru a minimiza impactul în situațiile de risc înalt.

Studiul actual a cercetat mai multe metode de clasificare a imaginilor cu scopul evaluării acurateței metodei în detecția pre-operatorie a potențialelor fistule pancreatice semnificative clinic.

Predicția a fost posibilă folosind modelul VGG 16, fără să sufere de overfitting. Studiul actual trece în evidență modele de antrenare bazate pe machine learning, dar suferă prin absența unui studiu de validare comparând cu scorurile tradiționale (fistula risk score și alternative fistula risk score).

Concluzii: Estimarea riscului de fistulă pancreatică semnificativă clinic beneficiază de implementări utilizând machine learning, dar necesită validare pe loturi mari și raportat la clasele de risc estimate tradițional prin clase de risc.

Declarația de conflict de interese.

Autorii declară că nu se află în conflict de interese financiare sau non financiare pentru datele și informațiile prezentate în articol. A fost obținut consimțământul informat scris de la pacientă pentru publicarea acestui raport de caz și a imaginilor însoțitoare

Bibliografie selectivă

1. Richter A, Niedergethmann M, Sturm JW, Lorenz D, Post S, Trede M. Long-term results of partial pancreaticoduodenectomy for ductal adenocarcinoma of the pancreatic head: 25-Year experience. World J Surg. 2003;
2. Trede M, Schwall G, Saeger HD. Survival

after pancreatoduodenectomy: 118 Consecutive resections without an operative mortality. Ann Surg. 1990;

3. Adam U, Makowiec F, Riediger H, Schareck WD, Benz S, Hopt UT. Risk factors for complications after pancreatic head resection. Am J Surg. 2004;
4. DeOliveira ML, Winter JM, Schafer M, Cunningham SC, Cameron JL, Yeo CJ, et al. Assessment of complications after pancreatic surgery: A novel grading system applied to 633 patients undergoing pancreaticoduodenectomy. Ann Surg. 2006;
5. Ecker BL, McMillan MT, Allegrini V, Bassi C, Beane JD, Beckman RM, et al. Risk Factors and Mitigation Strategies for Pancreatic Fistula after Distal Pancreatectomy: Analysis of 2026 Resections from the International, Multi-institutional Distal Pancreatectomy Study Group. Ann Surg. 2019;
6. Vin Y, Sima CS, Getrajdman GI, Brown KT, Covey A, Brennan MF, et al. Management and Outcomes of Postpancreatectomy Fistula, Leak, and Abscess: Results of 908 Patients Resected at a Single Institution Between 2000 and 2005. J Am Coll Surg. 2008;
7. Tien YW, Lee PH, Yang CY, Ho MC, Chiu YF. Risk factors of massive bleeding related to pancreatic leak after pancreaticoduodenectomy. J Am Coll Surg. 2005;
8. Gouma DJ, Van Geenen RCI, Van Gulik TM, De Haan RJ, De Wit LT, Busch ORC, et al. Rates of complications and death after pancreaticoduodenectomy: Risk factors and the impact of hospital volume. Ann Surg. 2000;
9. Murakami Y, Uemura K, Hayashidani Y, Sudo T, Hashimoto Y, Nakagawa N, et al. No mortality after 150 consecutive pancreatoduodenctomies with duct-to-mucosa pancreaticogastrostomy. J Surg Oncol. 2008;
10. Poon RTP, Fan ST, Lo CM, Ng KK, Yuen WK, Yeung C, et al. EXternal drainage of pancreatic duct with a stent to reduce leakage rate of pancreaticojejunostomy after pancreaticoduodenectomy: A prospective randomized trial. Ann Surg. 2007;
11. Pratt WB, Maithel SK, Vanounou T, Huang ZS, Callery MP, Vollmer CM. Clinical and economic validation of the international study group of pancreatic fistula (ISGPF) classification scheme. Ann Surg. 2007;
12. Hashimoto Y, Traverso LW. Incidence of pancreatic anastomotic failure and delayed gastric emptying after pancreatoduodenectomy in 507 consecutive patients: Use of a web-based calculator to improve homogeneity of definition. Surgery. 2010;
13. Lin JW, Cameron JL, Yeo CJ, Riall TS, Lillemoe KD. Risk factors and outcomes in postpancreaticoduodenectomy pancreaticocutaneous fistula. J Gastrointest Surg. 2004;
14. Yeo CJ, Cameron JL, Maher MM, Sauter PK, Zahurak ML, Talamini MA, et al. A prospective randomized trial of pancreaticogastrostomy versus pancreaticojejunostomy after pancreaticoduodenectomy. In: Annals of Surgery. 1995.
15. Chaudhuri S, Chatterjee S, Katz N, Nelson M, Goldbaum M. Detection of blood vessels in retinal images using two-dimensional matched filters. IEEE Trans Med Imaging. 1989;
16. Hoover A, Goldbaum M. Locating the optic nerve in a retinal image using the fuzzy convergence of the blood vessels. IEEE Trans Med Imaging. 2003;
17. Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. Commun ACM. 2017;
18. Zeiler MD, Fergus R. Visualizing and understanding convolutional networks. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). 2014.
19. Donahue J, Jia Y, Vinyals O, Hoffman J,

Zhang N, Tzeng E, et al. DeCAF: A deep convolutional activation feature for generic visual recognition. In: 31st International Conference on Machine Learning, ICML 2014. 2014.

20. Razavian AS, Azizpour H, Sullivan J, Carlsson S. CNN features off-the-shelf: An astounding baseline for recognition. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops. 2014.
21. Yosinski J, Clune J, Bengio Y, Lipson H. How transferable are features in deep neural networks? In: Advances in Neural Information Processing Systems. 2014.
22. Callery, M. P., Pratt, W. B., Kent, T. S., Chaikof, E. L., & Vollmer, C. M. (2013). A prospectively validated clinical risk score accurately predicts pancreatic fistula after pancreatoduodenectomy. *Journal of the American College of Surgeons*.
23. Mungroop, T. H., van Rijssen, L. B., van Klaveren, D., Smits, F. J., van Woerden, V., Linnemann, R. J., de Pastena, M., Klompmaker, S., Marchegiani, G., Ecker, B. L.,et al.. (2019). Alternative Fistula Risk Score for Pancreatoduodenectomy (a-FRS): Design and International External Validation. *Annals of Surgery*.
24. Mungroop, T. H., Klompmaker, S., Wellner, U. F., Steyerberg, E. W., Coratti, A., D'Hondt, M., de Pastena, M., Dokmak, S., Khatov, I., Saint-Marc, et al. (2021). Updated Alternative Fistula Risk Score (ua-FRS) to Include Minimally Invasive Pancreatoduodenectomy: Pan-European Validation. *Annals of Surgery*.
25. Schuh, F., Mihaljevic, A. L., Probst, P., Trudeau, M. T., Müller, P. C., Marchegiani, G., Besselink, M. G., Uzunoglu, F., Izbicki, J. R., Falconi, M., Fernandez-del Castillo, C., et al. (2021). A Simple Classification Of Pancreatic Duct Size and Texture Predicts Postoperative Pancreatic Fistula. *Annals of Surgery*.
26. Honselmann, K. C., Antoine, C., Frohneberg, L., Deichmann, S., Bolm, L., Braun, R., Lapshyn, H., Petrova, E., Keck, T., Wellner, U., & Bausch, D. (2021). A simple nomogram for early postoperative risk prediction of clinically relevant pancreatic fistula after pancreatoduodenectomy. *Langenbeck's Archives of Surgery*.